WHAT IS CLAIMED IS:

1. A method of classifying multivalued data stored in data storage means of a

computer system in a descending hierarchy, each datum being associated with

particular initial values of attributes that are common to the data, the method

5   comprising recursive steps of subdividing data sets, and wherein, during each step of

subdividing a set, discrete values are calculated for the attributes from the particular

initial values of the data attributes of said set, and wherein said set is subdivided into

subsets as a function of a homogeneity criterion calculated on the basis of the discrete

values for the attributes of said set.

10

2. A method of classifying data in a descending hierarchy according to claim 1,

wherein during the step of calculation of discrete values for the attributes, each initial

attribute is transformed into a discrete attribute.

15   3. A method of classifying data in a descending hierarchy according to claim 1,

wherein, during each step of subdividing a set, binary attribute values are calculated

from the particular initial attribute values of the data of said set, and wherein said set

is subdivided into subsets as a function of the binary values.

20   4. A method of classifying data in a descending hierarchy according to claim 3,

wherein, during the step of calculating the binary values for the attributes, for each

attribute that is numerical, the median value of the particular initial values of said

attribute in the data of said set is estimated and in that the value "true" is given to the binary attribute corresponding to said attribute for a datum of said set if the particular initial value of the numerical attribute of said datum is less than or equal to the estimated median value, else the value "false" is given thereto.

5

5. A method of classifying data in a descending hierarchy according to claim 4, wherein the estimated median value of a numerical attribute is obtained as follows:

- extracting extreme values from the set of values taken by the numerical attribute for the data of said set;

10   - calculating the mean of the remaining values; and

- allocating the value of said mean as the estimated median value.

6. A method of classifying data in a descending hierarchy according to claim 3, wherein, during the step of calculating the binary values for the attributes, for each

15   attribute that is symbolic, the modal value of the particular initial values of said attribute in the data of said set is estimated, and wherein the value "true" is allocated to the binary attribute corresponding to said attribute for a datum of said set is estimated if the initial particular value of the symbolic attribute of said datum is equal to the estimated modal value, else the value "false" is given thereto.

20

7. A method of classifying data in a descending hierarchy according to claim 6, wherein the modal value of a symbolic attribute is estimated as follows:

- the symbolic values taken by the data of said set for the symbolic attribute
are read;

- while reading the symbolic values, the first $\underline{m}$ different symbolic values
taken by the data of said set for the symbolic attribute are stored, where $\underline{m}$ is a

5   predetermined number;

- the symbolic value that appears most frequently is retained, amongst said $\underline{m}$
first different symbolic values; and

- the retained symbolic value is used as the estimate of the modal value.

10   8. A classification method according to claim 1, wherein said set is subdivided on the

basis of the discrete values of the most discriminating attribute, i.e. the attribute for

which a homogeneity criterion for all of the discrete values of the other attributes in

the resulting subsets is optimized.

15   9. A classification method according to claim 8, wherein, for any attribute, the

homogeneity criterion is an estimate of the expectation of the conditional probabilities

for correctly predicting the other attributes, given knowledge of this attribute.

10. A classification method according to claim 8, wherein, for certain attributes

20   marked a priori as being "taboo" by means of a particular parameter, the attribute

considered as being the most discriminating is the attribute that is not marked as being

taboo for which the homogeneity criterion for all of the discrete values of the other

attributes in the resulting subsets is optimized.